

Twitter as a Structured Information System

Dorja Mučnjak

Faculty of Humanities and Social Sciences Library

Ivana Lučića 3, HR-10000 Zagreb, Croatia

dmucnjak@ffzg.hr

Aleksandra Pikić

Faculty of Humanities and Social Sciences Library

Ivana Lučića 3, HR-10000 Zagreb, Croatia

apikic@ffzg.hr

Summary

In the past few years Twitter has achieved a tremendous success. Twitter is an online microblogging service, the tweets up to 140 characters allow people to use Twitter via SMS platforms or computers. With more than 200 million users and 155 million posted messages a day, Twitter has become a vast storehouse of content, i.e. information and knowledge. Such large amount of data that can be accessed on Twitter requires a structure in order to help users find information they need.

The purpose of this paper is to explore ways of structuring information on social networks, especially on Twitter. The aim of this study is to search for and present the most popular methods of gathering and structuring data published on the social networking site Twitter.

Key words: information and knowledge management, Twitter, hashtags, Twitter aggregators, Twitter directories

Introduction

In the past few years Twitter has achieved a tremendous success. As one of the three most popular social networking sites (Facebook, Twitter, LinkedIn), it has evidently found the path to its users. Twitter is an online microblogging service, the tweets up to 140 characters allow people to use Twitter via SMS platforms on most mobile phones, but the tweets can also be posted from computers – one of the reasons Twitter is so popular and widely used. Since the 140-character tweets do not allow extensive posts (especially in comparison to blogs where posts can be much longer), the users quickly realized that Twitter could be used as an ideal pointing device to other websites, for sharing links to longer articles, blog posts, audio-video materials, etc. Websites that were once dominantly accessed from Google search results are now seeing a growing number of new

visitors coming from shared links on Facebook or Twitter. In short, Twitter is a social networking site, a live searching tool and a link sharing place.

With more than 200 million users and 155 million posted tweets per day, Twitter has become a vast storehouse of content, i.e. information and knowledge. Such large amount of data that can be accessed on Twitter requires a structure in order to help users find information they need.

The purpose of this paper is to explore ways of structuring information on social networks, especially on Twitter. The aim of this study is to research and present the most popular methods of gathering and structuring data published on the social networking site Twitter.

Information and knowledge management

Today's society is a knowledge society. New data are being generated in a vast amount on a daily basis.

In order to transform this large amount of data into knowledge, first we have to contextualize the data, i.e. transform them into information and then, through the process of description and evaluation, transform the information into knowledge.¹

The knowledge management is the most important means of prosperity. In the knowledge society, information and knowledge management, as a concept, cannot be limited only to business intelligence gathering, but these processes should be integrated into the (re)production of everyday knowledge.²

In the last few decades, the information and communication technology (ICT) has been a great help in the knowledge management, i.e. in dissemination and (re)generation of knowledge. However, it must be said that ICT is a very useful tool, yet not the resource itself. The dissemination and generation of knowledge is primarily an interaction between people, and ICT is a means of transfer of knowledge.³ But, integration of ICT in everyday life has evidently changed the ways of communication, generation and dissemination of knowledge. For example, many people see Google as a unique source of information, it has become most visited "public library" without librarians as intermediaries who have skills and knowledge of retrieving, gathering, evaluating, and disseminating information.

At this point the educational process is no longer a process where a teacher transfers his knowledge to students. Today's education⁴ is a process of

¹ Afrić, Vjekoslav; Lasić-Lazić, Jadranka; Banek Zorica, Mihaela. Znanje, učenje i upravljanje znanjem. // Odabrana poglavlja iz organizacije znanja / Lasić-Lazić, J. (ed.). Zagreb : Zavod za informacijske studije, 2004; 37.

² Ibidem; 57.

³ Ibidem; 56.

⁴ Here, the term education is used for formal, non-formal and informal learning.

knowledge (co)production, i.e. a teacher and his students together produce the knowledge through a research process.⁵ Apart from teachers, a community also plays a significant role in the production of knowledge. The knowledge production and organization within community of peers has no longer a top-down, but a bottom-up structure.⁶ Such a structure is the result of collaboration and communication. A social network of peers is a productive generator of information, but a key question is how to organize information to make it usable and actual knowledge.⁷

Social networking sites – communication and collaboration tools or sources of information?

Everyday technological changes and progress modify the production and transfer of knowledge. In the past few years online social networking sites (SNS) have facilitated communication within social networks. Social networking sites are low-cost easy-to-use tools that allow people to share and use a lot of information quickly and easily.

As communication and collaboration tools, these sites are structured as personal (or egocentric)⁸ with an individual in the center of their own community; they are organized primarily around people not around interests. That means that one is a “friend” or a “follower” of a person one share interests with – the foundation of this relationship is trust, this is the point of difference at forums that are organized around interests and where trust is not that important because participants are strangers. In that sense, social networking sites are great communication and collaboration tools that allow people to interact with the element of stronger confidence in data.

As said before, social networks in the real world play a significant role in the knowledge (co)production – participants share information with friends, colleagues, relatives... In virtual online world the situation is the same, but facilitated with tools (as social networking sites) that allow people to create larger communities. Yes, social networks in the real world create a source of information, but a mass interaction between people on social networking sites that

⁵ Afrić, Vjekoslav. Upravljanje sadržajem učenja i znanja. // Odabrana poglavlja iz organizacije znanja / Lasić-Lazić, J. (ed.). Zagreb : Zavod za informacijske studije, 2004; 63-64.

⁶ Avram, Gabriela. At the Crossroads of Knowledge Management and Social Software. // The Electronic Journal of Knowledge Management. 4(2006), 1; 1.
<http://www.ejkm.com/issue/download.html?idIssue=8> (20.5.2011.)

⁷ Upravljanje sadržajem učenja i znanja. 2004; 64.

⁸ Boyd, Danah M.; Ellison, Nicole B. Social network sites : definition, history, and scholarship. // Journal of Computer-Mediated Communication. 13(2008), 1; 219.
<http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html> (27.5.2011.)

produce a vast amount of data makes an even more considerable source of information.

To cope with this great amount of data, one needs guidelines. Depres i Chauvel⁹ proposed a framework for categorizing five types of knowledge management activities that could be easily used when retrieving information on social networking sites:

- **Scan/map** - pointing to the world of overview of data;
- **Acquire/capture/create** – associated with the world of research, development and creation;
- **Package/codification/representation/storing** – related to the world of databases, information and knowledge bases, organizational memory;
- **Apply/share/transfer** – related to the world of competencies, teamwork, intranets and cross border sharing;
- **Reuse/innovate/evolve/transform** – associated to the world of leverage, intellectual assets and innovation.

Aforementioned activities are mainly meant for organization of business intelligence, but these principles could be used for organization of information of any kind and in any environment. Evidently, social networking sites could be used in any of these five processes:

Scan/map – social networking sites are an ideal place to provide information on interesting people, data, projects, etc. For example, following people or organizations gives us an overview of information they share with their friends/followers.

Acquire/capture/create – browsing social networking sites, e.g. Twitter, is mainly focused on discovering the needs, expertise and offers of other people and indirectly of other organizations.

Package/codification/representation/storing – social networking sites, as a storage of content, enable to find, collect and store information immediately and, not less important, without intermediaries.

Apply/share/transfer – on social networking sites, one can easily share information.

Reuse/innovate/evolve/transform – information posted on social networking sites remain available for a longer period of time (if not forever) and they could be meant for everybody (in case of public access) or only for members of a particular group (in case of limited access).¹⁰

In the beginning social networking sites were primarily communication and collaboration tools. However, different features that social networking sites offer to users (publishing videos, photos, other multimedia, etc.) have transformed them into a significant source of content.

⁹ At the Crossroads of Knowledge Management and Social Software. 2006; 5.

¹⁰ Ibidem; 5-7.

But, it has also become evident that the great amount of information generated on social networking sites needs some kind of structuralization in order to facilitate retrieving information.

Three Twitter's characteristics

In this paper, we speak mainly about Twitter – the online microblogging service with more than 200 million registered users who post more than 155 million *tweets* (140-character posts) a day.¹¹

After a big success of Facebook, the concept of Twitter as a microblogging service could have sounded weird. Facebook, apart from being a social networking tool, offers interesting multimedia features, such as posting photographs, creating photo-albums, posting videos, playing games or using different applications designed for fun. In addition, on Facebook one can update his/her status, i.e. one posts a note about himself/herself, about what he/she is doing, what he/she is thinking about... Unlike Facebook, Twitter allows posting only short notes, called *tweets*.

Twitter was originally created as a communication platform that allows posting short messages (*tweets*) as an answer to the question "*What are you doing?*", just like updating the status on Facebook (but limited to 140 characters). Without the possibility to post other interesting objects (photos, videos, music...), Twitter users had soon discovered its simplicity and straight-forward nature and came up with an idea how to make it interesting. Suddenly, it became a great tool to say what was happening around you, to comment on daily news, to inform others about important discoveries... Since the length of tweets is very limited, people started to use Twitter dominantly as a **link-sharing tool**, as a pointing device to other objects – websites, discussions, blogs, etc. To shorten long URLs, people started to use online URL shortening services (for example: <http://bit.ly>, <http://tinyurl.com>). Short URLs allow otherwise long web addresses to be referred to in a Tweet. The company itself recognized the way users had used Twitter, so, a few years after its launch, they decided to replace the familiar "*What are you doing?*" tag line that had sat atop status update box with a tag line "*What's happening?*"

Created primarily as a **social networking site**, Twitter allows users to *follow* (to receive other people's tweets on the profile) other users (people, institutions, firms...) as *followers*. Users can choose whether they want their accounts to be public or protected. In case of a public account, profile pages would be visible to everyone. A protected account allows only approved followers to view that

¹¹ What is Twitter? <http://business.twitter.com/basics/what-is-twitter> (25.5.2011.)

account's tweets. The average Twitter user follows a few dozens of people: a collage of friends, colleagues, and a handful of celebrities.¹²

Twitter has always been very easy to use not only on computers but also via SMS platform. Twitter itself is a very simple program. However, it allows third party programmers to develop different useful applications – for computer and for mobile device use (especially BlackBerry and iPhone). Such applications make Twitter more user friendly, contributing to its popularity.

Evidently, Twitter is a well-used communication tool, or social networking device. But the number of 155 million of Tweets a day shows that Twitter is also a great source of content. It has become important information storage and knowledge generator.

Twitter has another interesting feature: a search box that gives a **real-time** view onto the conversation of just about any topic imaginable. On the profile page, Tweets are being updated in a real time. This makes Twitter an ideal search engine for the newest events. It is also the reason why a certain number of websites, once dominantly accessed from Google search results, now see a growing number of new visitors coming from shared links on Twitter.¹³

To use Twitter as a source of information more easily, it is necessary to structuralize its content. There are a number of community conventions (without any intervention in program) for grouping topics, events or people by certain symbols.

One of the most used community conventions is a *hashtag* (#), the symbol used to mark relevant keywords or topics in a Tweet, i.e. to categorize and join all tweets with the same topic and show them easily in Twitter Search. If one uses a hashtag on a public account, anyone who does a search for that hashtag may find it.

Another common convention for organizing Tweets is the symbol @ (*at*). It can be used to directly address other user (convention *@Reply*) – @username in the beginning of a Tweet; or to mention other user (convention *@Mention*) – @username anywhere in a Tweet.

These conventions are not always sufficient for structuralizing content. Fortunately, there is a great number of third party applications that can help retrieve information more successfully. The most popular ways of structuralizing content published on Twitter are Twitter aggregators and directories.

¹² Johnson, Steven. How Twitter Will Change The Way We Live (in 140 characters or less). // Time. 173(15.6.2009), 23; 34-35. <http://www.time.com/time/business/article/0,8599,1902604,00.html> (20.5.2011.)

¹³ Ibidem.

Social network aggregation

Social network aggregation is a process of collecting content from multiple social networking services, such as Twitter. The aggregation tasks are most frequently performed by social network aggregators, the tools that pull together information into a single location or help a user to consolidate multiple social networking profiles into a single one. Various aggregation services provide tools or widgets which allow users to collect messages, track friends, gather bookmarks, search across multiple social networking sites, read rss feeds for various sites, access their profiles from a single location, etc.¹⁴

Another way of collecting tweets of various users is through link directories. Web directories or link directories list links to other websites organized by category and subcategory. Most of the directories are general, but some niche directories focus on restricted regions, single languages, or specialist sectors. They can also be used for collecting profiles of various Twitter users or trending topics on Twitter.

The topic of this paper are aggregators that collect tweets of various Twitter users in a single location and link directories of Twitter profiles or topics. The focus of our interest will not be on aggregators that consolidate multiple social networking profiles into a single one, but on the aggregators that pull together Tweets into a single location.

In the previous sections we have presented the necessity to structure the content on Twitter.

Apart from the Twitter community conventions of # (hashtags) and @ (Mentions) used in tweets, the aggregation of content is another way of structuralizing information on Twitter. Twitter allows third party programmers to create different applications for Twitter, and many aggregation tools are made by third parties. So, it must be pointed out that the content published on Twitter is mainly not structuralized by Twitter tools but by third party applications or community conventions.

Twitter aggregators and directories

Since Twitter is a rather new social networking tool, there is only a small number of scholarly articles or other literature on the subject. However, there are some guidelines that can help find information on Twitter content aggregation such as blogs,^{15, 16} directories,¹⁷ and web portals.¹⁸

¹⁴ Social Network Aggregation. Interpiseo. URL: http://www.interpiseo.com/resources/general_info_articles/Social%20Network%20Aggregation.pdf (20.5.2011.)

¹⁵ Odden, L. 12 Twitter Stream Aggregators To Make You Smarter. / TopRank : Online Marketing Blog. URL: <http://www.toprankblog.com/2009/02/12-twitter-stream-aggregators/> (20.5.2011.)

In terms of aggregation of content published on Twitter, there are several ways of collecting information: directories, aggregators, popularity counters, and newspapers. Some of these directories or aggregators are capable of processing both functions at the same time: listing and aggregating.

Twitter directories are lists of links to Twitter accounts organized by certain criteria: by people (very common among politicians) or by subject (list of Twitter profiles posting tweets about the same subject). There are also directories that collect links to other Twitter directories or aggregators.

Examples:

blog.govtwit – list of Twitter aggregators on politics

Just Tweet It – organized by subject

U.S. Air Force – list of Twitter profiles on U.S. Air Force

Twitter directories/aggregators – as mentioned earlier, some of these applications have two functions – collecting links to individual Twitter profiles and aggregating tweets posted on those profiles. Many politicians have discovered these tools and use them frequently.

Examples:

GovLuv – U.S. government representatives

politter.com – Japanese politicians

poliTwitter.ca – Canadian politicians

politwitter.de – German politicians

TweetCommons – Canadian politicians

TweetCongress – U.S. congressmen

Twitter aggregators automatically harvest tweets on different subjects from different Twitter accounts.

Examples:

Europatweets – aggregator of tweets on EU

Retailer Twitter Aggragation – aggregator of tweets on retail

Retweetist – aggregator of retweets

¹⁶ Worth, Jon. Who are the main players in the EU-twitthersphere? / Jon Worth Blog. 27.01.2010. <http://www.jonworth.eu/who-are-the-main-players-in-the-eu-twitthersphere/> (25.5.2011.)

¹⁷ Lunceford, Steve. 50+ government Twitter and social media directories. / Blog.govtwit. 25.10.2010. <http://www.blog.govtwit.com/2010/08/25/50-government-twitter-and-social-media-directories/> (20.5.2011.)

¹⁸ Finn, Greg. 3 Aggregation Tools For Tracking Tweets. 21.9.2009. <http://searchengineland.com/tracking-tweets-how-to-twitter-aggregation-tools-22699> (20.5.2011.)

Tweetmeme – aggregator of tweets by subject

Twistori – aggregator of tweets containing one of the following words: love, hate, think, believe, feel, wish

TwitLinks – aggregator of tweets on technologies

Popularity counters

Twitter **popularity counters** aggregate the most popular tweets.

Examples:

TwitterCounter

Twiturls

Twiturlly

Twopular

Newspapers

The most interesting case of aggregation is the site that organizes links shared on Twitter into an easy newspapers-style format, and can be created for any Twitter user, list or hashtag (#).

Examples:

paper.li

Conclusion

Twitter is a well-used link-sharing and live-searching social networking site. With more than 155 million of tweets a day it has become a great source of information. Twitter is a very simple tool which does not offer possibility of structuralizing this vast amount of data, but it allows users and third parties to create different applications that can help structuralize the content.

The most popular ways of structuralizing the content on Twitter are the community conventions: the hashtags (#) and the @ symbol, which pull together all the tweets with these symbols. The hashtags mark specific topics and the @ symbol helps to gather all the tweets a person was mentioned in.

Twitter has a very simple and straight-forward nature; it does not offer many different ways of structuralizing information but it allows others to create interesting applications. Many of its users took this opportunity and have created many different applications that harvest information from Twitter and structuralize the content. Most common applications are directories (lists of links organized by some criteria), aggregators (tools that pull together information into a single location), popularity counters (organized by criteria of popularity of tweets), and newspaper-style sites (sites which organize links shared on Twitter).

All these useful applications and conventions have grown out of the necessity of structuralizing the great amount of information published on Twitter.

References

- Afrić, Vjekoslav. Upravljanje sadržajem učenja i znanja. // Odabrana poglavlja iz organizacije znanja / Lasić-Lazić, J. (ed.). Zagreb : Zavod za informacijske studije, 2004; 62-101.
- Afrić, Vjekoslav; Lasić-Lazić, Jadranka; Banek Zorica, Mihaela. Znanje, učenje i upravljanje znanjem. // Odabrana poglavlja iz organizacije znanja / Lasić-Lazić, J. (ed.). Zagreb : Zavod za informacijske studije, 2004; 33-61.
- Avram, Gabriela. At the Crossroads of Knowledge Management and Social Software. // The Electronic Journal of Knowledge Management. 4(2006), 1; 1-10.
<http://www.ejkm.com/issue/download.html?idIssue=8> (20.5.2011.)
- Boyd, Danah M.; Ellison, Nicole B. Social network sites : definition, history, and scholarship. // Journal of Computer-Mediated Communication. 13(2008), 1; 210-230.
<http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html> (27.5.2011.)
- Finn, Greg. 3 Aggregation Tools For Tracking Tweets. 21.9.2009.
<http://searchengineland.com/tracking-tweets-how-to-twitter-aggregation-tools-22699> (20.5.2011.)
- Johnson, Steven. How Twitter Will Change The Way We Live (in 140 characters or less). // Time. 173(15.6.2009), 23; 32-37.
<http://www.time.com/time/business/article/0,8599,1902604,00.html> (20.5.2011.)
- Lunceford, Steve. 50+ government Twitter and social media directories. / Blog.govtwit. 25.10.2010. <http://www.blog.govtwit.com/2010/08/25/50-government-twitter-and-social-media-directories/> (20.5.2011.)
- Odden, Lee. 12 Twitter Stream Aggregators To Make You Smarter. TopRank : Online Marketing Blog. <http://www.toprankblog.com/2009/02/12-twitter-stream-aggregators/> (20.5.2011.)
- Social Network Aggregation. Interpiseo.
http://www.interpiseo.com/resources/general_info_articles/Social%20Network%20Aggregation.pdf (20.5.2011.)
- What is Twitter? <http://business.twitter.com/basics/what-is-twitter> (25.5.2011.)
- Worth, Jon. Who are the main players in the EU-twitersphere? / Jon Worth Blog. 27.01.2010.
<http://www.jonworth.eu/who-are-the-main-players-in-the-eu-twitersphere/> (25.5.2011.)

Hyperlinks

- blog.govtwit <http://www.blog.govtwit.com> (30.5.2011.)
- Europatweets. <http://www.europatweets.eu/> (30.5.2011)
- GovLuv <http://govluv.org/> (29.5.2011.)
- Just tweet it <http://justtweetit.com/> (30.5.2011.)
- Paper.li <http://paper.li/> (30.5.2011.)
- Politter.com <http://politter.com/en/> (30.5.2011.)
- poliTwitter.ca <http://politwitter.ca/> (30.5.2011.)
- Politwitter.de <http://politwitter.de/> (30.5.2011.)
- Retailer Twitter Aggregation <http://www.noturnonred.org/twitter/> (30.5.2011.)
- Retweetist <http://retweetist.com/> (30.5.2011.)
- Social Media. U.S. Air Force. <http://www.af.mil/socialmedia.asp> (30.5.2011.)
- Tweermeme <http://tweermeme.com/> (30.5.2011.)
- TweetCommons <http://ca.tweetcommons.com/polis/running> (30.5.2011.)
- TweetCongress http://tweetcongress.org/channels/view/us_media (30.5.2011.)
- Twistori <http://twistori.com/> (30.5.2011.)
- TwitLinks <http://twitlinks.com/> (30.5.2011.)
- Twitt(url)y <http://twitturlly.com/> (30.5.2011.)
- Twitturls <http://twitturls.com/> (30.5.2011.)
- Twopular <http://twopular.com/> (30.5.2011.)
- TwitterCounter <http://twittercounter.com/> (28.5.2011.)